

Classification of Breast Mass Abnormalities using Denseness and Architectural Distortion

S. Baeg^{*} and N. Kehtarnavaz⁺

^{*} Texas Instruments, Inc., Dallas, 12500 TI Boulevard, MS8723, Dallas, TX 75243, USA

⁺ Department of Electrical Engineering, EC33, University of Texas at Dallas, 2601 N. Floyd Rd.
Richardson, TX 75083-0688, USA

Received 18 March 2002; accepted 25 June 2002

Abstract

This paper presents an electronic second opinion system for the classification of mass abnormalities in mammograms into benign and malignant categories. This system is designed to help radiologists to reduce the number of benign breast cancer biopsies. Once a mass abnormality is detected and marked on a mammogram by a radiologist, two textural features, named denseness and architectural distortion, are extracted from the marked area. The denseness feature provides a measure of radiographic denseness of the marked area, whereas the architectural distortion feature provides a measure of its irregularity. These features are then fed into a neural network classifier. Receiver operating characteristic (ROC) analysis was conducted to evaluate the system performance. The area under the ROC curve reached 0.90 for the DDSM database consisting of 404 biopsy proven masses. A sensitivity analysis was also performed to examine the robustness of the introduced texture features to variations in sizes of abnormality markings.

Keywords: Classification of mass abnormality, mammography, texture features, electronic second opinion, denseness and architectural distortion.

1. Introduction

Breast cancer is a major cause of death in women. Nonetheless, it has been shown that the death rate from breast cancer can be reduced by the early detection of malignancies. X-ray mammography is the most widely available screening procedure for the early detection of breast cancer [1]. After performing a mammography exam, patients are often asked to do a routine screening, short interval follow-up, or biopsy examination depending on the findings.

After detecting breast abnormalities on mammograms, radiologists assess their likelihood of malignancy or benignity. Currently, not only patients having highly suggestive of malignancy category undergo surgical biopsies, but also patients having probable malignancy category are sent for biopsies. This is primarily due to the difficulty associated with distinguishing between benign and malignant abnormalities by visual inspection. A considerable number of biopsies yield benign results. When a biopsy is recommended, the likelihood of a benign outcome is 70% to 80% [2].

This research was funded by the Texas Instruments sponsored DSP Program at Texas A&M University.
Correspondence email kehtar@utdallas.edu

Recommended for acceptance by Maria Petrou
ELCVIA ISSN: 1577-5097
Published by Computer Vision Center / Universitat Autonoma de Barcelona, Barcelona, Spain

As stated in [3], "Many women who undergo a surgical biopsy suffer from adverse psychological consequences even after they receive a benign result." Therefore, a reduction in the number of benign biopsies, while achieving a desired sensitivity, will not only lower the health care costs associated with biopsies but also lessen patient suffering caused by undergoing the traumatic experience of surgical biopsy.

Many efforts have been made on computer analysis of mammograms. Several computer-aided diagnosis (CAD) systems have been developed for the detection and classification of two major signs of malignancy in mammograms: namely clustered microcalcifications and masses. A considerable amount of work has been done on the detection of microcalcifications, e.g. [4], [5], [6]. The focus of this work is on the classification of masses. There have been studies on the classification of masses using various image features such as shape, fractal dimension, and laws texture features, e.g. [7], [8], [9], [10], [11], [12].

Polakowski [7] used shape, size, contrast, and Laws texture features for the classification of masses. A neural network classifier was trained on 33 malignant, 39 benign and 45 other false masses, correctly recalling all 33 malignant cases. Brzakovic [8] proposed a classifier that employed four features: size, shape, edge distance variation, and edge intensity variation. Out of a total of 25 cases, 85% of them were classified correctly. Matsubara [12] used fractal dimension features for the classification of masses reporting a perfect classification rate when tested on 13 cases: 7 benign masses and 6 malignant masses. Rangayyan [9] utilized the "acutance" feature to discriminate between benign and malignant tumors. The acutance feature measures the averaged gray-level difference between the inside and outside of a tumor boundary. The tumor boundary was traced manually. The classification rate achieved was 92.6% using 54 tumors (16 circumscribed benign, 7 circumscribed malignant, 12 spiculated benign, and 19 spiculated malignant lesions). El-Faramawy [10] used various shape features including compactness, Fourier descriptors, moments, and chord-length statistics. He examined different combinations of these features on a database of 54 tumors reporting a classification rate of 76%. Menut [11] proposed a parabolic modeling method to characterize a tumor's boundary. The classification of 54 masses resulted in a classification rate of 76%.

Our hypothesis in this work was that texture features, based on radiographic denseness and irregularity, would be more effective when examining a wide variety of cases or a large database. Most breast cancers cause high x-ray attenuation because they are radiographically dense. On the contrary, benign masses cause relatively low x-ray attenuation and are intermingled with fat that is radiographically less dense. In addition, malignant masses exhibit irregularity in shapes and margins, whereas benign masses show more homogeneity.

Therefore, in this paper, we have developed two new texture features to capture the denseness and irregularity of a marked mass abnormality area. These texture features are then used as the input to a neural network classifier. Section II discusses the introduced texture features. The neural network classifier is covered in Section III. The analysis and classification of mammograms are then presented in Section IV.

2. Developed Texture Features

Unfortunately, there is no universally agreed upon definition of texture. Various mathematical tools have been used to define texture features. Co-occurrence matrix [13], fractal [14], wavelet [15], and granulometry [16], [17] based methods have been used to define texture features. Using such texture features, we found that there is a fair amount of overlap between the distributions of benign and malignant categories for a relatively large database considered in this work. Figure 1 illustrates two examples of such texture features based on co-occurrence matrix and fractals. Malignant cases are represented by plus signs "+" and benign cases by circle signs "o". As can be seen from this figure, there exists a considerable overlap between the feature values for benign and malignant cases. In what follows, we present two new texture features, named denseness and architectural distortion, that are specifically designed to have relatively non-overlapping distributions for malignant and benign mass abnormalities.

2.1. Denseness Texture Feature

The denseness texture feature is derived from differences in the radiographic denseness between benign and malignant masses. As Kopans points out in *Breast Imaging* [2] "Most breast cancers are radiographically very dense for their size and high in x-ray attenuation. They are generally not intermingled with fat, and their

x-ray attenuation increases toward their centers." Since high x-ray attenuation produces high intensity values in mammograms, the center of a tumor usually appears to be bright. On the contrary, he states "Most benign lesions are radiographically less dense and low in x-ray attenuation. Most benign masses appear as encapsulated radiolucent masses in mammograms. They are usually intermingled with fat with relatively low x-ray attenuation."

Based on the above clinical difference between benign and malignant lesions, an image denseness feature D is devised here by going through a series of transformations of a mammogram J as follows:

$$J \xrightarrow{\Psi_x} B \xrightarrow{\Psi_m} E \xrightarrow{V} D,$$

where the transformation Ψ_x generates a binary image B consisting of local maxima/minima along the rows and columns of a marked mass abnormality area, and the transformation Ψ_m generates a morphologically eroded image E . The function V returns the number of remaining pixels in the image E .

In order to locate the local maxima/minima, the method introduced by Rosenfeld [18] is utilized here. Let $f(r, c)$ represent image intensity at row r and column c . A pixel at (r_0, c_0) is a local row maximum if

$$f(r_0, c_0 - 1) < f(r_0, c_0) > f(r_0, c_0 + 1).$$

That is, a pixel becomes a local row maximum if its intensity level is higher than that of its neighboring row pixels. If the intensity is constant in an interval $\{(r_0, c) | i \leq c \leq j\}$ such that

$$f(r_0, i - 1) < f(r_0, c) > f(r_0, j + 1),$$

then the pixel at the center of the interval is taken to be the local row maximum. Local column maximum, row minimum and column minimum are obtained in a similar manner. The set of maxima/minima M is considered to be the collection of all local row maximum, row minimum, column maximum, and column minimum points. Figure 2 illustrates the local column maxima found for two different synthetic images: J_A and J_B . The image J_A has high intensity levels increasing toward its center while the intensity levels in J_B fluctuate over the image.

The maxima/minima points are then used to construct a binary image B as follows:

$$B(r, c) = \begin{cases} 0, & \text{if } (r, c) \in M \\ 1, & \text{if } (r, c) \notin M \end{cases}.$$

Although the population of maxima/minima in the binary image B can be used as a measure of denseness, population alone does not capture the spatial variation associated with maxima/minima and hence does not adequately describe textural denseness. In other words, mammograms with the same population can possess different spatial variations of maxima/minima.

In order to incorporate the spatial variation of maxima/minima into the denseness feature, the morphological erosion operator Ψ_m [19] is applied to the binary image B . Let S denote a horizontal or vertical structuring element. The morphological erosion is

$$B \ominus S = \{z | S_z \subseteq B\}$$

where S_z denotes the translation of the structuring element S to z . Erosion generates all the locations where the structuring element is contained in the foreground pixels of B . Since the foreground pixels in B are the non-maxima/minima points, erosion generates locally dense regions corresponding to the structuring

element. Now if the spatial variation of maxima/minima is different, then the erosion set $B\Theta S$ will be different even though the populations may be the same. As a result, this way both the population and the spatial variation of maxima/minima contribute to the denseness of a marked mass abnormality area.

Next a binary image E is constructed from the eroded image as follows:

$$E(r, c) = \begin{cases} 1, & \text{if } (r, c) \in B\Theta S \\ 0, & \text{if } (r, c) \notin B\Theta S \end{cases}$$

Considering that the eroded locations correspond to the non-maxima/minima points, the number of elements in the eroded set $B\Theta S$ is equal to the number of locally dense regions. Let E_h and E_v indicate the eroded images by a horizontal and a vertical structuring element, respectively. The denseness texture feature D is defined as:

$$D = \sum_{\forall(r,c) \in E_h} E(r, c) + \sum_{\forall(r,c) \in E_v} E(r, c) \quad (1)$$

Equation (1) indicates the denseness feature is proportional to the number of eroded locations. Basically, the denseness feature becomes high if the marked area contains many locally dense regions. A denseness feature value close to zero indicates the marked area has a few locally dense regions, while a value far off zero indicates it has many locally dense regions.

2.2 Appropriate Size of Structuring Element

It should be noted that the denseness feature varies depending on the size of the structuring element. This size is determined here by maximizing the area under the ROC curve [20], [21]. In essence, the ROC curve is used to provide a measure of class separability for the denseness feature whose probability density function is unknown. The ROC curve gives the relationship between true-positive fraction (TPF) and false-positive fraction (FPF) for various decision thresholds τ . Let x denote an observation of the denseness feature. TPF and FPF are defined as follows:

$$\begin{aligned} TPF(\tau) &= \frac{\#\{x \mid x > \tau, \forall x \in X_M\}}{\#\{x \mid \forall x \in X_M\}}, \\ FPF(\tau) &= \frac{\#\{x \mid x > \tau, \forall x \in X_B\}}{\#\{x \mid \forall x \in X_B\}}, \end{aligned}$$

where X_M and X_B are the set of feature values extracted from the malignant and benign mass abnormality areas, respectively, and $\#$ is the number of elements.

In order to find an appropriate size, many ROC curves are generated using different sizes $\gamma = (\gamma_h, \gamma_v)$ of the horizontal and vertical structuring elements. The size γ_o that maximizes the ROC curve area A_z is then selected to be an appropriate size of the 2D structuring element.

2.3. Architectural Distortion Texture Feature

The architectural distortion texture feature is derived from gradient differences between malignant and benign masses. In *Breast Imaging* [2], Kopans states that "Breast cancer does not always produce a mammographically visible mass, but it frequently disrupts the normal tissues in which it develops. The probability of malignancy increases as a lesion becomes more irregular in shape. On the contrary, benign lesions are not architecturally distorted but have round, oval, or lobulated shape. Since the margins are not irregular in most benign lesions, benign masses appear encapsulated."

Based on such a clinical irregularity difference, the following image texture feature is devised here to quantify irregularities in intensity levels:

$$AD = \frac{\sum_{\forall(r,c) \in J} [w_1 * |G_x(r,c)|]^3}{N_1} + \frac{\sum_{\forall(r,c) \in J} [w_2 * |G_y(r,c)|]^3}{N_2} \quad (2)$$

where

$$w_1 = \frac{f(r,c) + f(r,c+1)}{2}, \quad w_2 = \frac{f(r,c) + f(r+1,c)}{2},$$

and G_x denotes gradient in the horizontal direction and G_y gradient in the vertical direction, which are computed as follows [19]:

$$\begin{aligned} G_x(r,c) &= f(r,c+1) - f(r,c) \\ G_y(r,c) &= f(r+1,c) - f(r,c) \end{aligned}$$

The introduced architectural distortion texture feature reflects the sum of weighted gradients. In (2), the denominators N_1 and N_2 are

$$\begin{aligned} N_1 &= \#\{(r,c) \in J \mid G_x(r,c) \neq 0\} \\ N_2 &= \#\{(r,c) \in J \mid G_y(r,c) \neq 0\} \end{aligned}$$

Each term in (2) is normalized by N_1 and N_2 instead of the total number of pixels in order to make the architectural distortion feature less correlated with the denseness feature. When normalized by the total number of pixels, some locally dense regions heavily influence this feature because such regions produce zero gradients leading to a small numerator in (2). The division by N_1 and N_2 excludes the zero-gradient pixels. The power 3 of gradient in (2) is found to make the architectural distortion feature less dependent on the denseness property, since it yields a much larger value for severe intensity changes than for mild changes. The weights in (2) account for the global denseness of masses.

In essence, if a marked mass abnormality area is heterogeneous, i.e. neighboring pixels having noticeable intensity differences, the gradients tend to be large leading to a large architectural distortion feature value. If a marked mass abnormality area is homogeneous, i.e. neighboring pixels having similar intensity values, gradients tend to be small leading to a small architectural distortion feature value. The texture heterogeneity or homogeneity is hence characterized by the strength of the AD feature. This feature does not depend on any user-defined parameter.

Figure 3 illustrates two synthetic images: J_C and J_D . The intensity levels in J_C change significantly between nearby pixels, whereas the intensify levels in J_D change mildly between nearby pixels. Figure 4 shows the profile of weighted gradients $w_2 * |G_y(r,c)|$ for each column in J_C and J_D . The image J_C produces larger gradient values than the image J_D due to its higher heterogeneity.

3. Mass Abnormality Classification

A three-layer backpropagation neural network [22] was chosen as the classifier after examining several parametric and non-parametric classifiers. The neural network classifier was found to produce the highest classification rate as compared to the other classifiers. The neural network used consisted of two input nodes, thirty hidden nodes and one output node. The number of hidden layer nodes was determined by experimentation to generate the smallest mean-square-error. The denseness and architectural distortion

texture features extracted from a marked mass abnormality area were fed into the input nodes. The network was then trained to produce the output value of 1 if the input feature vector belonged to a malignant case, and the output value of 0 if it belonged to a benign case.

Basically, the classification process was done as follows: First, the denseness and architectural distortion texture features were extracted from the area marked by the radiologist to be a mass abnormality area. Then, these features were fed into the neural network classifier to provide a malignancy or benignity value.

3.1. ROC Analysis

As with all lesion classification problems, the likelihood that disease is present corresponds to a probability estimate. Different radiologists prefer different decision thresholds τ . In this computer-aided system, the classifier yields a single number η for each input mass image, which is then compared to different thresholds τ for classification.

Given the pathology of masses, TPF and FPF were easily determined. TPF is the number of malignant masses correctly classified as malignant divided by the total number of malignant masses. FPF is the number of benign masses wrongly classified as malignant divided by the total number of benign masses. The area under ROC curve A_z was then used as a measure of classification performance.

4. Experimental Results

4.1. Mammogram Database

The mammograms used in our experiments consisted of the DDSM (Digital Database for Screening Mammography) database at the University of South Florida [23]. Each mammogram is associated with an overlay file, which gives information about the abnormality area marked by a radiologist. These marked masses were classified in our experiments. A total of 404 biopsy proven mass images (218 malignant and 186 benign) having a resolution of 50 μm were made available to us at the time of this experiment to carry out the following analysis.

The DDSM database also provides the breast density information for each mammogram. There are four ratings specified for breast density: fat containing (density 1), low density (density 2), equal density (density 3), and high density (density 4). Out of the 404 masses, 127 masses are specified to have density 1, 184 density 2, 85 density 3, and 8 density 4. As per their shapes, 48 masses are indicated to possess round, 115 oval, 102 lobulated, 114 irregular, and 25 distorted shapes. As per their margins, 149 masses are indicated to possess circumscribed, 71 microlobulated, 85 ill-defined, 21 obscured, and 78 spiculated margins.

4.2. Denseness Feature Results

The scatter or distribution of the denseness feature for the database of 404 mass images is shown in Figure 5. Malignant cases are represented by plus signs "+" and benign cases by circle signs "o". Two typical mass images J_1 and J_2 are shown in Figure 6. The overlaid contour is the marking made by the radiologist. The mass image J_1 shown in Figure 6(a) is a malignant mass with a lobulated shape and ill-defined margin. The mass image J_2 shown in Figure 6(b) has an oval shape and a circumscribed margin.

Figure 7 displays the morphologically eroded images E_1 and E_2 . These erosions were performed by a horizontal structuring element of length 5. The image E_1 shows a large area of dense region in the middle. The high x-ray attenuation of the breast cancer in J_1 causes this outcome.

4.3. Selection of Appropriate Size of Structuring Element

Figure 8 shows the 3D plot of the class separability measure and its gradients for 49 different sizes of the structuring element. As can be seen, the class separability increases as the sizes approached 1. Also, the

gradients increased as the sizes approached 1. The class separability reached a maximum of 0.86 when the length of the horizontal structuring element was 1 and the length of the vertical structuring element was 5. Therefore, the combination $\gamma_o = (1,5)$ was selected as an appropriate size of the structuring element and used in the subsequent analysis.

4.4. Architectural Distortion Feature Results

The scatter or distribution of the architectural distortion feature for the database of 404 mass images is shown in Figure 9. Malignant cases are represented by plus signs “+” and benign cases by circle signs “o”. Two typical masses J_3 and J_4 are shown in Figure 10. The mass shown in Figure 10(a) is a malignant mass with an architecturally distorted shape and an ill-defined margin. The mass shown in Figure 10(b) is a benign mass with a lobulated shape and a circumscribed margin.

Figure 11 shows the profile of weighted gradients $w_2^* |G_y(r,c)|$ for each column. The malignant mass J_3 exhibit larger weighted gradient values than the benign mass J_4 . When a mass shape is architecturally distorted, neighboring pixels possess noticeable intensity differences causing larger weighted gradients. The ill-defined margin also contributes to the larger gradients. On the contrary, lobulated shapes or circumscribed margins, in general, possess lower variations of intensity. In Figure 11, the weighted gradients are relatively low in columns from 1 to 200. The homogeneous regions in the top left-hand corner produce these lower values.

Figure 12 shows the scatter plot of the denseness and architectural distortion texture features. The correlation coefficient between the denseness and architectural distortion features was found to be 0.12. As can be seen from this figure, the samples are randomly scattered showing little correlation between the two introduced texture features.

4.5. Classification Performance

Out of a total of 404 masses, 80 masses (40 malignant and 40 benign) were randomly selected and used to train the neural network classifier. The remaining 324 masses were used to test the classification system. It should be emphasized that we only used one fifth of the data for training. Normally, one half of the data is used for training.

A total of 120 different neural network designs were considered. The classification performance or the area underneath the ROC curve for the test data was obtained to be 0.902 on average using both denseness and architectural distortion features. The 95% confidence interval for the ROC area was found to be (0.899 - 0.904) with standard deviation of 0.014. Some randomly chosen ROC curves are shown in Figure 13. In order to plot the ROC curves, a total of 100 threshold values were considered with the step size of 0.01.

The above classification performance was compared with those of the commonly used texture features. The ROC areas were found to be 0.61, 0.65, 0.68, and 0.81 for granulometry, co-occurrence, fractal, and wavelet-based texture features, respectively. This was another indication that the introduced texture features were more effective than the previously used texture features for the classification of mass abnormalities.

4.6 Sensitivity to Marking Size

As depicted in Figure 14, the size of an abnormality marking can vary from radiologist to radiologist. Since the features are extracted from the pixels of a marked abnormality area, the feature values slightly vary depending on the radiologist who performs the marking. We carried out a sensitivity analysis to evaluate the robustness of the developed texture features to varying sizes of markings.

It is expected that the radii of markings by most radiologists for the same masses to be close. So, for analysis purposes, it is reasonable to assume that the radii of the markings are normally distributed with the mean value of r_o . Let v denote the deviation or noise added to the nominal marked radius r_o . Let us assume this noise is normally distributed with zero mean and standard deviation σ_N . This standard deviation

indicates the amount of deviation from the nominal marking. The radius r' of a typical marking is thus derived as follows:

$$r' = r_o + v.$$

Three levels of σ_N were examined, $\sigma_N = 5, 10, 20$. One hundred ROC curves were generated for one hundred statistical simulations of the three noise levels. It was obtained that the maximum change in the ROC area was within 2% of the original area. In addition, a hypothesis test was performed to see if the performance of $\sigma_N = 5$ was statistically different from $\sigma_N = 20$. Let μ_1 and μ_2 denote the mean area A_z for these two cases. According to the hypothesis test:

$$\begin{aligned} H_0 : \mu_1 - \mu_2 &= 0 \quad (\text{that is, no difference in mean } A_z) \\ H_a : \mu_1 - \mu_2 &\neq 0 \end{aligned},$$

there was statistically no difference in the mean performance between these cases at the significance level of 0.19.

5. Conclusion

This paper has addressed the classification of mass abnormalities in mammograms into benign and malignant categories. Considering that the distributions of the commonly used texture features were found to be overlapping for our relatively large database of biopsy proven masses, we devised two new texture features to capture denseness and architectural distortion differences between benign and malignant masses. These features correlate well with similar clinical differences. The denseness feature was designed to capture radiographic denseness of breast cancer, while the architectural distortion feature to capture its irregularity. The area under ROC curve was obtained to be 0.90 using a database of 404 biopsy proven masses. The classification performance was not statistically sensitive to the variation of markings with the significance value of 0.19.

As depicted in Figure 15, all the developed software algorithms have been incorporated into a graphical user interface (GUI) to allow our electronic second opinion system to be easily used in clinical environments without the need for the user to know the details of the algorithms. This GUI reads a mammogram, extracts the denseness and architectural distortion features from a marked abnormality area, and launches the classifier. This system can be used as an electronic second opinion to lower the number of benign biopsies, in particular for cases when its outcome agrees with the radiologist's assessment. Interested readers can obtain this GUI from <http://www.utdallas.edu/~nxk019000/eso.html>.

References

- [1] C. D'Orsi and D. Kopans, "Mammography interpretation: The BI-RADS method," *American Family Physician*, Vol. 55, No. 5, April 1997.
- [2] D. Kopans, *Breast Imaging*. New York: Lippincott-Raven Publishers, 1998.
- [3] J. Brett, J. Austoker, and G. Ong, "Do women who undergo further investigation for breast screening suffer adverse psychological consequences? A multi-center follow-up study comparing different breast screening result groups five months after their last breast screening appointment," *American Journal of Public Health*, vol. 20, no 4, pp. 396-403, 1998.
- [4] J. Dengler, S. Behrens, and J. F. Desaga, "Segmentation of microcalcifications in mammograms," *IEEE Transactions on Medical Imaging*, vol. 12, no. 4, pp. 634 - 642, 1993.
- [5] T. C. Wang, and N. B. Karayiannis, "Detection of microcalcifications in digital mammograms using wavelets," *IEEE Transactions on Medical Imaging*, vol. 17, no. 4, pp. 498 - 509, 1998.

- [6] R. N. Strickland, and H. I. Hahn, "Wavelet transforms for detecting microcalcifications in mammograms," *IEEE Transactions on Medical Imaging*, vol. 15, no. 2, pp. 218 - 229, 1996.
- [7] W. Polakowski, D. Cournoyer, S. Rogers, M. DeSimio, and D. Ruck, "Computer-aided breast cancer detection and diagnosis of masses using difference of Gaussians and derivative-based feature saliency," *IEEE Transactions on Medical Imaging*, vol. 16, no. 6, pp. 811 – 819, 1997.
- [8] D. Brzakovic, X. Luo, and P. Brzakovic, "An approach to automated detection of tumors in mammograms," *IEEE Transactions on Medical Imaging*, vol. 9, no. 3, pp. 233 – 241, 1990.
- [9] R. Rangayyan, N. El-Faramawy, and J. Desautels, O. Alim, "Measures of acutance and shape for classification of breast tumors," *IEEE Transactions on Medical Imaging*, vol. 16, no. 6, pp. 799 – 810, 1997.
- [10] N. El-Faramawy, R. Rangayyan, J. Desautels, and O. Alim, "Shape factors for analysis of breast tumors in mammograms," in *Proceedings Canadian Conference on Electrical and Computer Engineering*, Calgray, AB, Canada, May 1996, pp. 355 – 358.
- [11] O. Menut, R. Rangayyan, and J. Desautels, "Classification of breast tumors via parabolic modeling of their contours," in *Proceedings 1997 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, Victoria, BC, Canada, August 1997, pp. 1002 – 1005.
- [12] T. Matsubara, H. Fujita, S. Kasai, M. Goto, Y. Tani, T. Hara, and T. Endo, "Development of new schemes for detection and analysis of mammographic masses," in *Proceedings Intelligent Information Systems. IIS'97*, Grand Bahamas Island, Bahamas, December 1997, pp. 63-66.
- [13] R. Haralick, *Computer and Robot Vision*. Boston: Addison-Wesley, 1992.
- [14] S. Peleg, J. Naor, R. Hartley, and D. Anvir, "Multi-resolution texture analysis and classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 4, pp. 518-523, 1984.
- [15] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*. Upper Saddle River, New Jersey: Prentice Hall, 1998.
- [16] S. Baeg, A. T. Popov, V. Kamat, S. Batman, K. Sivakumar, N. Kehtarnavaz, E. Dougherty, and R. Shah, "Segmentation of mammograms into distinct morphological texture regions," in *Proceedings Eleventh IEEE Symposium on Computer Based Medical Systems*, Lubbock, Texas, 1998, pp. 20-25.
- [17] Y. Chen and E. R. Dougherty, "Gray-scale morphological granulometric texture classification," *Optical Engineering*, vol. 33, no. 8, pp. 2713-2722, 1994.
- [18] A. Rosenfeld and E. Troy, "Visual texture analysis," University of Maryland, College Park, Maryland, Technical Report 70-116, 1970.
- [19] R. Jain, R. Kasturi and B.G. Schunck, *Machine Vision*. New York: McGraw-Hill, Inc., 1995.
- [20] C. E. Metz, "ROC methodology in radiologic imaging," *Investigative Radiology*, vol. 21, no. 9, pp. 720-733, 1986.
- [21] C. W. Therrien, *Decision Estimation and Classification*. New York: John Wiley & Sons, 1989.
- [22] M. Hagan, H. Demuth, and M. Beale, *Neural Network Design*. Boston: PWS Publishing Company, 1996.
- [23] M. Heath, K. Bowyer, D. Kopans, P. Kegelmeyer, R. Moore, K. Chang, and S. Munishkumaran, "Current status of the Digital Database for Screening Mammography," in *Digital Mammography*, N. Karssemeijer, M. Thijssen, J. Hendriks, and L. V. Erning, Ed. Boston: Kluwer Academic Publishers, 1998, pp. 457-460.

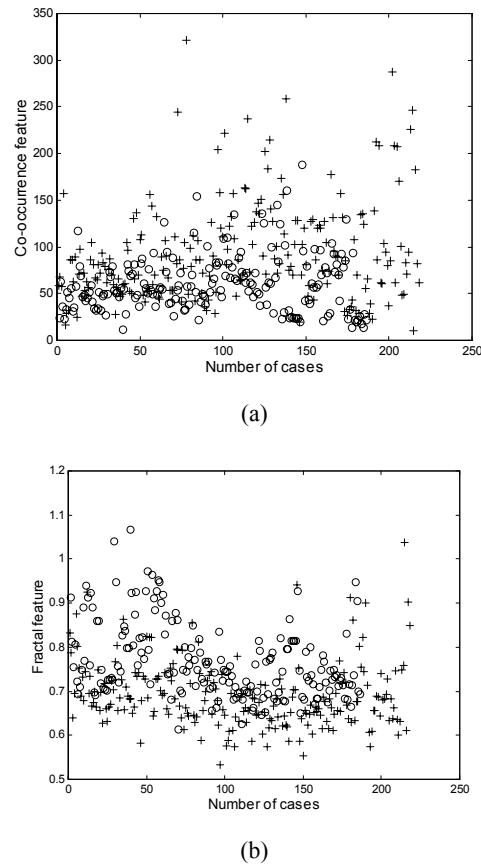


Figure 1. Scatter plot for feature values: (a) co-occurrence feature, (b) fractal feature. ‘+’s indicate values obtained for malignant masses and ‘o’s for benign masses. There exists too much overlap between the malignant and benign cases.

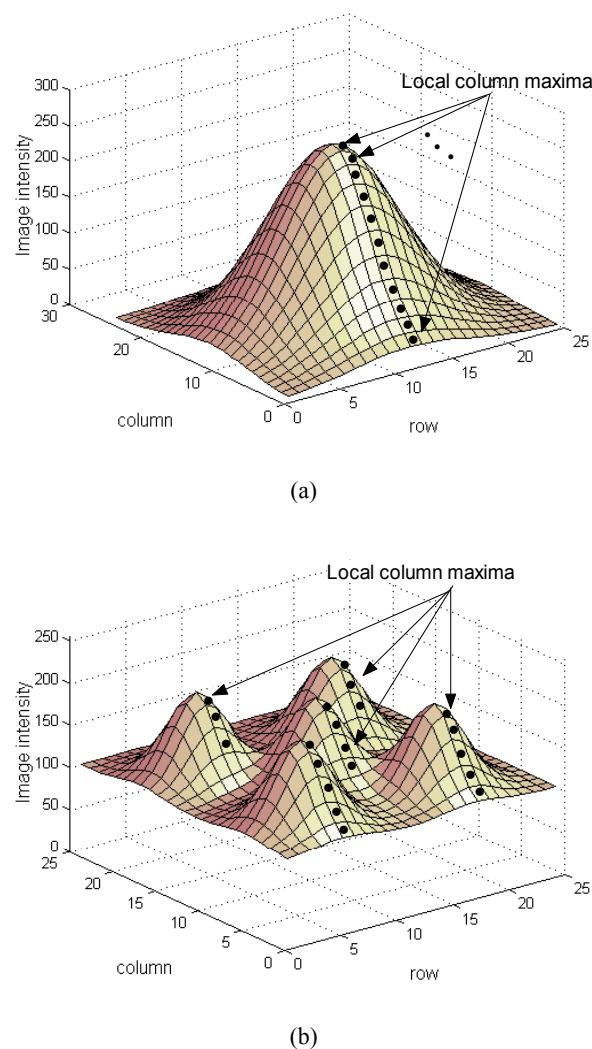


Figure 2. Synthetic images having different radiographic denseness: (a) an image J_A with high x-ray attenuation at the center, (b) an image J_B intermingled with low x-ray attenuation.

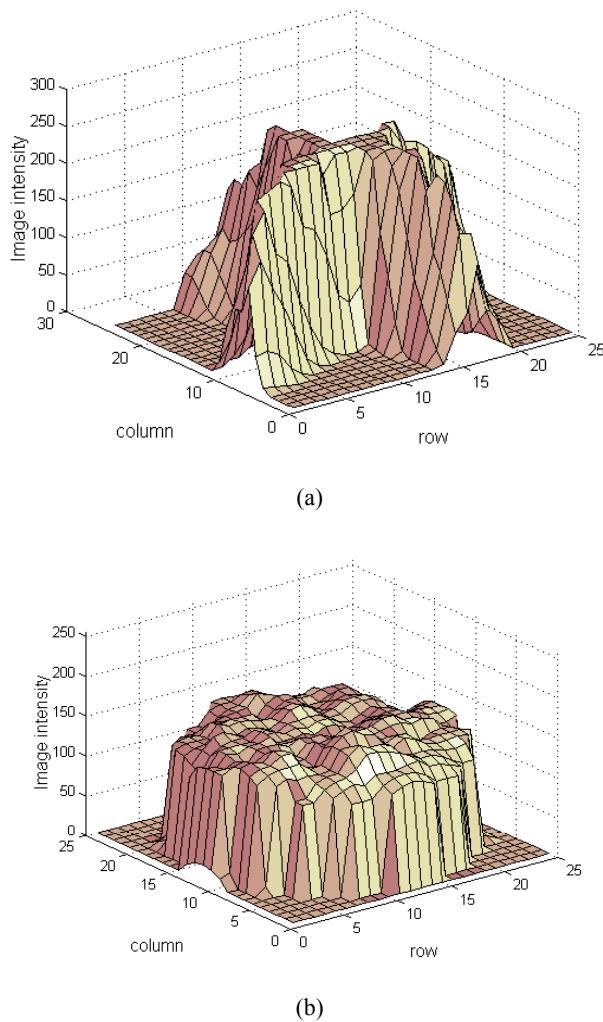


Figure 3. Synthetic images having different intensity variations: (a) an image J_C with severe intensity variations, (b) an image J_D with relatively mild intensity variations.

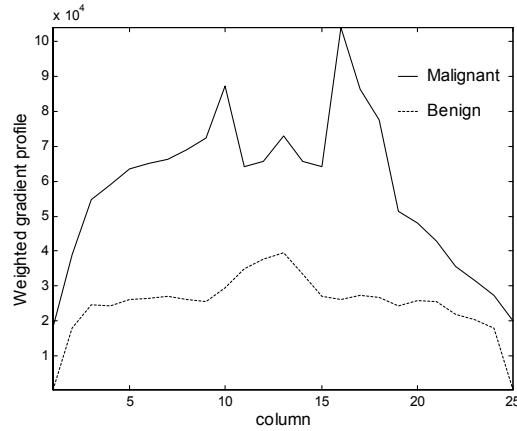


Figure 4. Column profile of weighted gradients $w_2 * |G_y(r, c)|$, malignant mass shows higher values than benign mass.

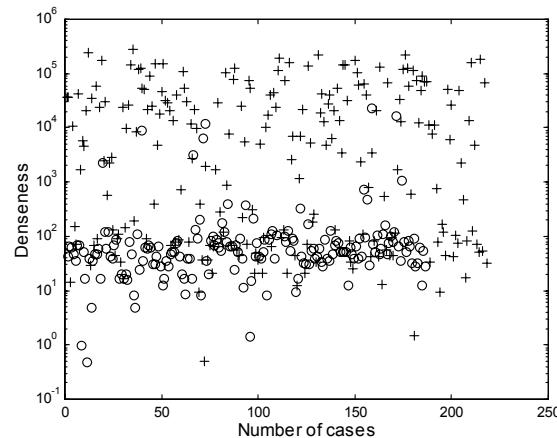
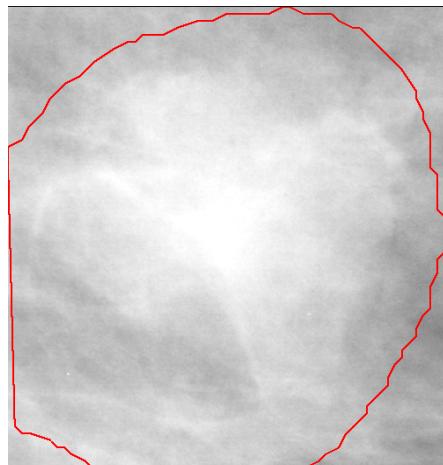
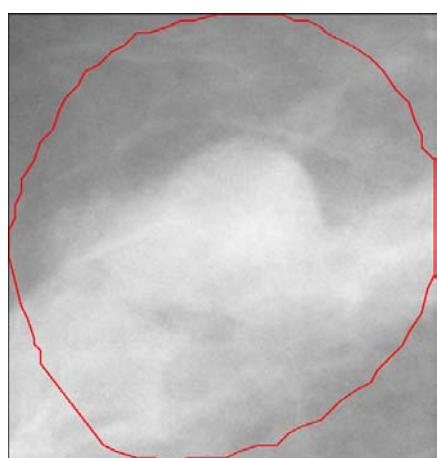


Figure 5. Scatter plot of denseness feature, '+'s indicate values obtained from malignant masses and 'o's from benign masses. In general, feature values from malignant cases have higher values than benign cases.

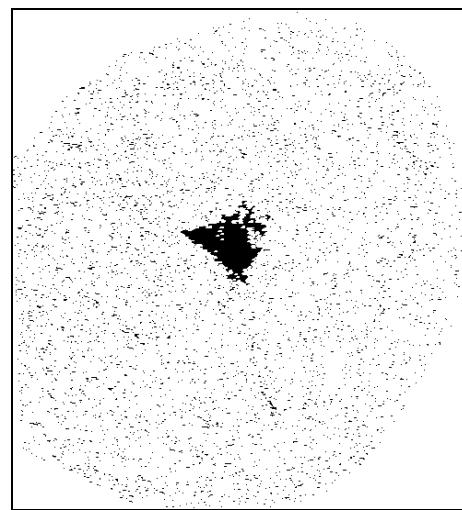


(a)

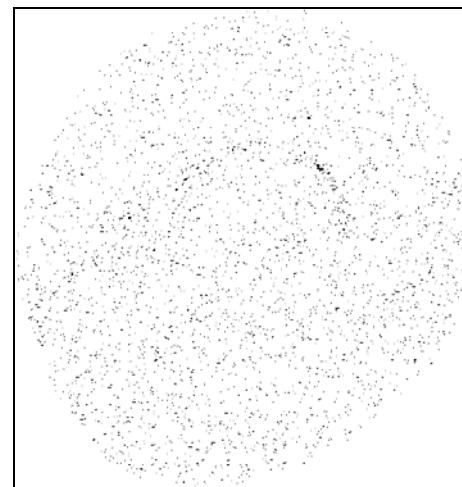


(b)

Figure 6. Typical mass images marked by a radiologist exhibiting different x-ray attenuation: (a) a malignant mass J_1 , and (b) a benign mass J_2 .



(a)



(b)

Figure 7. Morphologically eroded images: (a) eroded image E_1 from image J_1 , and (b) eroded image E_2 from image J_2 . Dark points represent eroded locations.

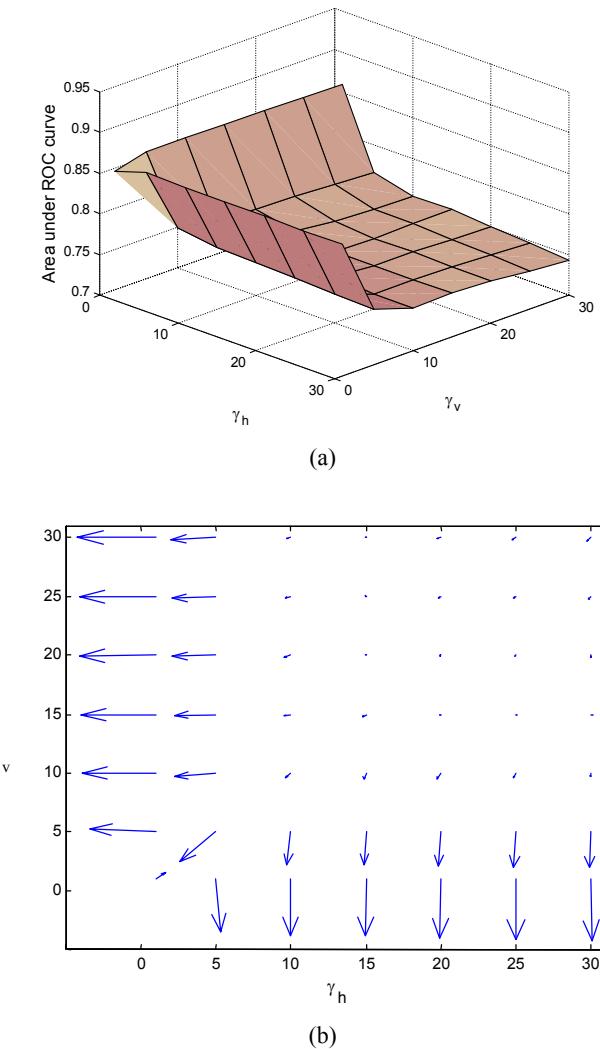


Figure 8. Class separability measure by ROC for different sizes of structuring elements: (a) ROC curve area, and (b) gradient of ROC curve area.

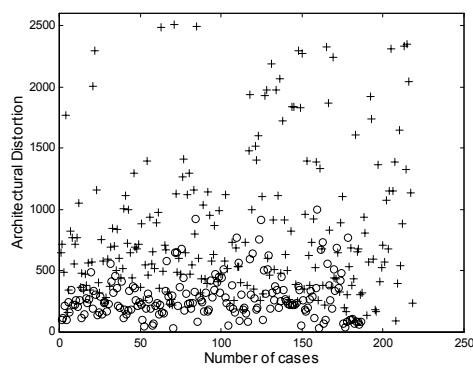


Figure 9. Scatter plot of architectural distortion feature. '+'s indicate values obtained from malignant masses and 'o's from benign masses.

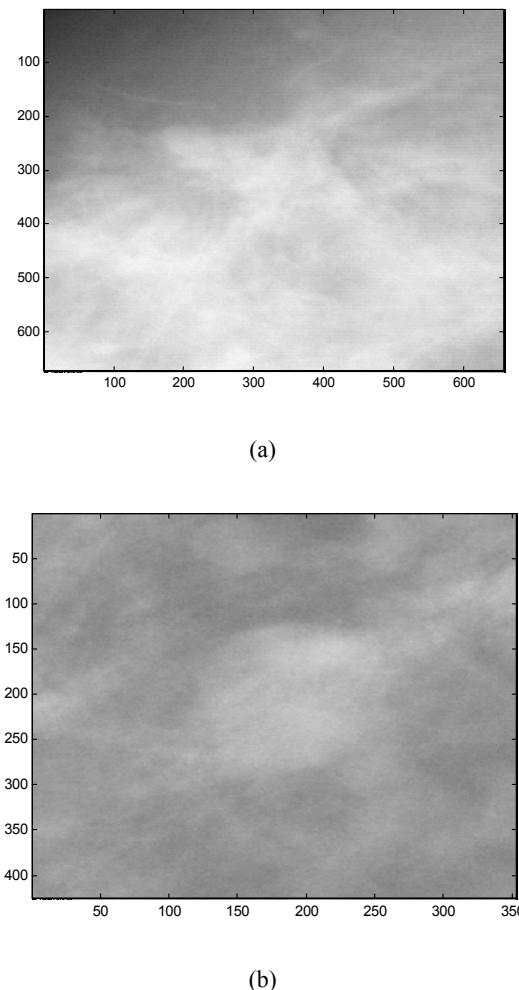


Figure 10. Typical malignant and benign masses: (a) a malignant mass J_3 with architectural distortion shape and ill-defined margin, and (b) a benign mass J_4 with lobulated shape and circumscribed margin.

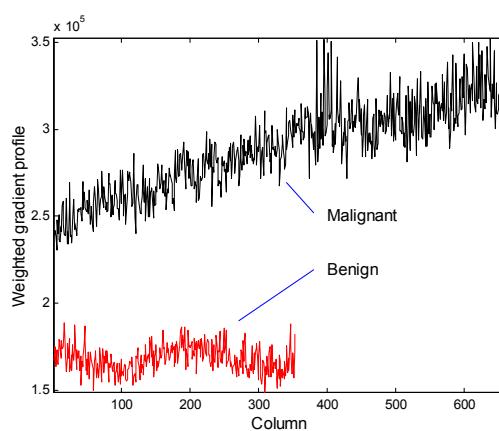


Figure 11. Column profile of weighted gradients $w_2 * |G_y(r,c)|$, malignant case shows higher values than benign case.

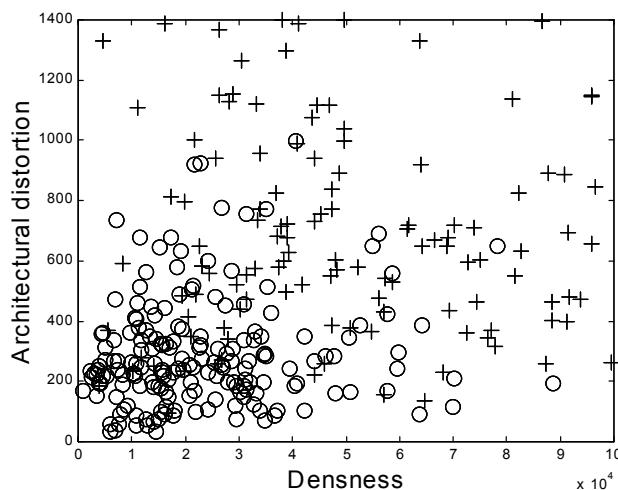
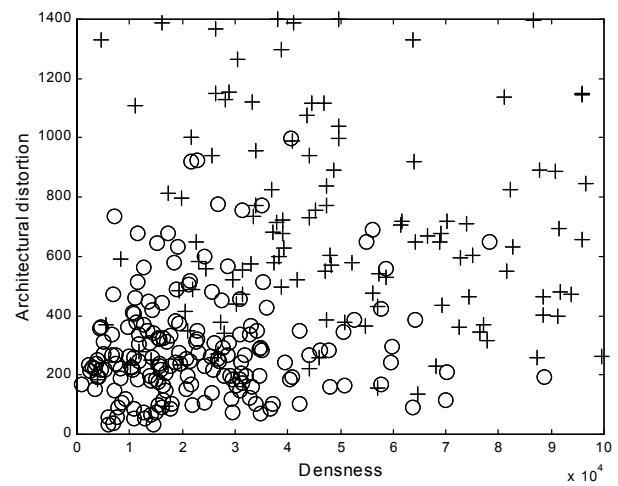


Figure 12. Scatter plot of the denseness feature and architectural distortion feature giving a correlation coefficient of 0.12.

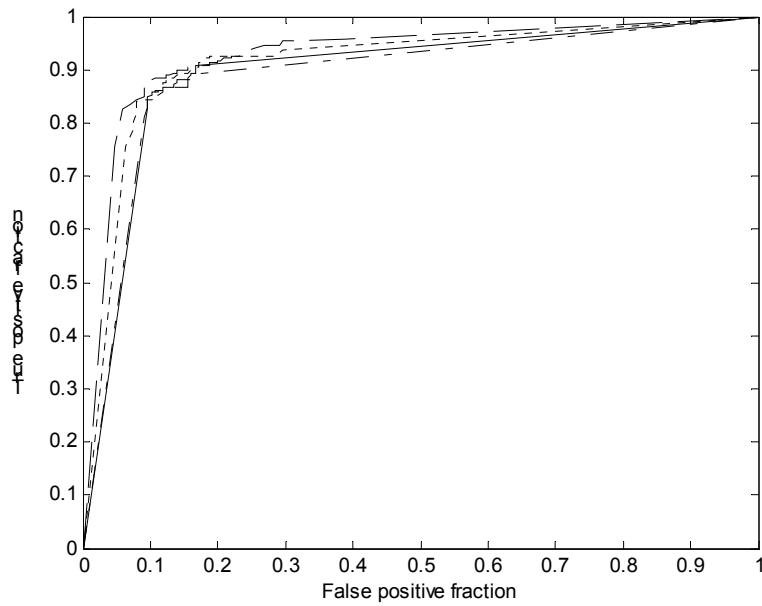


Figure 13. Randomly chosen ROC curves when using denseness and architectural distortion texture features, 120 different neural network designs were used generating an average ROC area of 0.902 with a 95% confidence interval of (0.899-0.904).

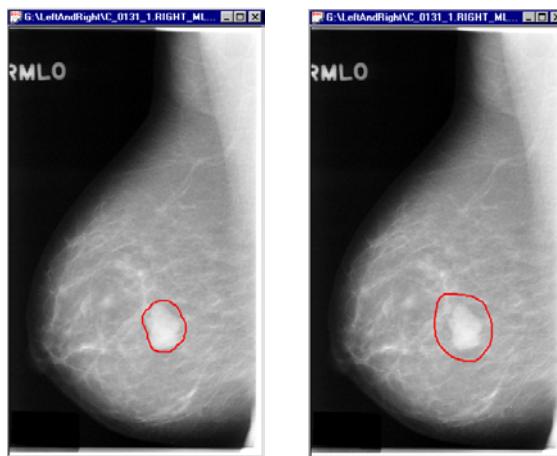


Figure 14. Different sizes of markings for the same mass considering that marking can vary from radiologist to radiologist.

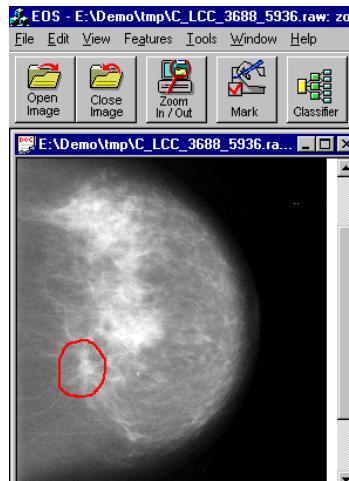


Figure 15. A snapshot of the electronic second opinion system. The classifier can be launched after marking a mass abnormality area.